Journal of
Marine Science and Engineering

MDPI

Article

Estimation of Artificial Reef Pose Based on Deep Learning

Yifan Song [1,2], Zuli Wu [1,*], Shengmao Zhang [1,*], Weimin Quan [1], Yongchuang Shi [1], Xinquan Xiong [1,3] and Penglong Li [1,3]

1 Key Laboratory of Fisheries Remote Sensing, Ministry of Agriculture and Rural Affairs, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China; syf_fyss@163.com (Y.S.); 15142892813@163.com (P.L.)
2 College of Information, Shanghai Ocean University, Shanghai 201306, China
3 School of Navigation and Naval Architecture, Dalian Ocean University, Dalian 116023, China
* Correspondence: wuzl@ecsf.ac.cn (Z.W.); zhangsm@ecsf.ac.cn (S.Z.)

Abstract: Artificial reefs are man-made structures submerged in the ocean, and the design of these structures plays a crucial role in determining their effectiveness. Precisely measuring the configuration of artificial reefs is vital for creating suitable habitats for marine organisms. This study presents a novel approach for automated detection of artificial reefs by recognizing their key features and key points. Two enhanced models, namely, YOLOv8n-PoseRFSA and YOLOv8n-PoseMSA, are introduced based on the YOLOv8n-Pose architecture. The YOLOv8n-PoseRFSA model exhibits a 2.3% increase in accuracy in pinpointing target key points compared to the baseline YOLOv8n-Pose model, showcasing notable enhancements in recall rate, mean average precision (mAP), and other evaluation metrics. In response to the demand for swift identification in mobile fishing scenarios, a YOLOv8n-PoseMSA model is proposed, leveraging MobileNetV3 to replace the backbone network structure. This model reduces the computational burden to 33% of the original model while preserving recognition accuracy and minimizing the accuracy drop. The methodology outlined in this research enables real-time monitoring of artificial reef deployments, allowing for the precise quantification of their structural characteristics, thereby significantly enhancing monitoring efficiency and convenience. By better assessing the layout of artificial reefs and their ecological impact, this approach offers valuable data support for the future planning and implementation of reef projects.

Keywords: artificial reefs; YOLOv8; key point detection; pose estimation

Citation: Song, Y.; Wu, Z.; Zhang, S.; Quan, W.; Shi, Y.; Xiong, X.; Li, P. Estimation of Artificial Reef Pose Based on Deep Learning. *J. Mar. Sci. Eng.* **2024**, *12*, 812. https://doi.org/10.3390/jmse12050812

Academic Editor: Francesco Tiralongo

Received: 15 April 2024
Revised: 3 May 2024
Accepted: 10 May 2024
Published: 13 May 2024

1. Introduction

The construction of artificial reefs aims to restore damaged coral reef ecosystems, protect and enhance fisheries resources, contribute to the restoration and improvement of marine ecological environments, and mitigate the impacts of human activities on the natural environment [1]. Recent studies have shown that artificial coral reefs exhibit numerous advantages. In 2020, da Silva G. V. et al. demonstrated that artificial reefs can effectively deflect longshore currents, alter sediment pathways, and thus induce long-term morphological changes [2]. Similarly, in 2020, Glarou M. et al. found that artificial coral reefs can significantly increase fish species diversity, enhance coral cover and biomass, and facilitate interactions among marine organisms, all with positive effects [3]. Moreover, in 2023, Firth L. B. et al. discovered that while diving resources hold economic value, excessive exploitation may harm ecosystems like coral reefs [4]. In addition to environmental benefits, artificial reefs offer direct economic advantages. Acarli D. et al. analyzed the species composition and distribution percentages of a lobster-designed artificial reef model in the Inderkum Bay area along the Turkish Mediterranean coast in 2020 [5]. Their results revealed that artificial coral reefs can enhance regional species diversity. Artificial reefs can also provide effects similar to those of natural coral reefs. Zhang R. et al. found in 2021 that artificial reef areas can support nutritional structures and pathways similar to those of natural reefs [6]. The most common designs for artificial reefs are cubes and complex structures made up of various shapes, mostly composed of concrete (79%), and placed in the ocean to create a coral reef environment [7]. The majority of artificial reefs are submerged at depths of less than 50 m, with most located between 10 and 20 m [8].

Studying and monitoring artificial coral reefs long-term is crucial to enhance our understanding of their impact on marine ecosystems and to promote sustainable marine ecological management [9]. In 2022, Lymperaki et al. emphasized the importance of conducting long-term research and monitoring on artificial coral reefs [10]. Furthermore, in 2021, Brochier et al. found that successful artificial coral reefs require consideration of factors such as local economic conditions, governance capacity, and illegal fishing, suggesting the establishment of long-term monitoring systems and adjustments in design and location for sustainable management [11]. Paxton et al. in 2022 further underscored the significance of artificial coral reefs in marine ecosystem restoration and biodiversity enhancement, proposing their inclusion in marine planning to drive sustainable marine ecological management [12]. In a study by Hemery et al. in 2022, various monitoring technologies were compared, with the conclusion that COTS 360-degree video cameras are suitable for monitoring artificial reefs (AR) and fish aggregating devices (FAD) efficacy in marine energy facilities [13].

The increasing recognition of the importance of coral reef ecosystems and their dependent ecosystems has led researchers to use deep learning technology for automatic annotation and monitoring of coral reefs. In recent years, various deep learning methods have been proposed and proven crucial in coral reef detection, classification, and conservation. Back in 2016, Mahmood et al. introduced a method using deep learning technology for automatic annotation of coral reefs, aiming to efficiently and accurately annotate new coral reef images by training neural networks to recognize different types of coral reefs [14]. In a recent review by Paxton et al. in 2020, the environmental threats faced by coral reefs, their significance, and the application of machine learning in conservation were discussed [15]. Gonzalez-Rivero et al. explored the role of artificial intelligence in coral reef monitoring in 2020, highlighting the improved accuracy and repeatability in estimating underwater biodiversity, along with the reduced time and cost required for data processing and reporting using deep learning convolutional neural networks (CNN) on coral reef datasets [16]. Xiong et al. designed a deep learning model based on Faster-RCNN and SSD to address the limitations of traditional algorithms in artificial reef recognition in 2021 [17]. Moreover, Sharan et al. presented an automatic coral classification system based on convolutional neural networks (CNN) and image enhancement in 2021 [18]. Experimental results demonstrate the effectiveness of the proposed method in learning and predicting accurately, showing potential for automatic coral classification that can be applied to large-scale datasets. In recent years, there has been a growing trend toward using deep learning technology for artificial reef detection. Deep learning not only facilitates efficient and precise annotation and monitoring of coral reefs but also plays a pivotal role in conservation and sustainable fisheries development [19]. This study presents an automated method for detecting and estimating the orientation of artificial reefs using deep learning technology. Through the use of advanced visual models, this method successfully recognizes artificial reefs, enhancing detection speed and accuracy. It provides technical support for marine biology research, conservation efforts, and sustainable fisheries development, thereby promoting the application and management of artificial reefs in marine ecosystems. The key innovations of this study are outlined below:

1. The introduction of a novel method based on sonar images for artificial reef recognition. By identifying key points to determine the posture of artificial reefs, this study proposes an innovative approach for estimating reef posture.
2. The construction of a dataset of artificial reefs based on Oculus sonar is the focus of this study.

3. Adoption of the YOLO v8n-pose framework [20] in the model design, incorporating various techniques to enhance model performance, accuracy, and efficiency while also considering constraints related to model size and computational power.

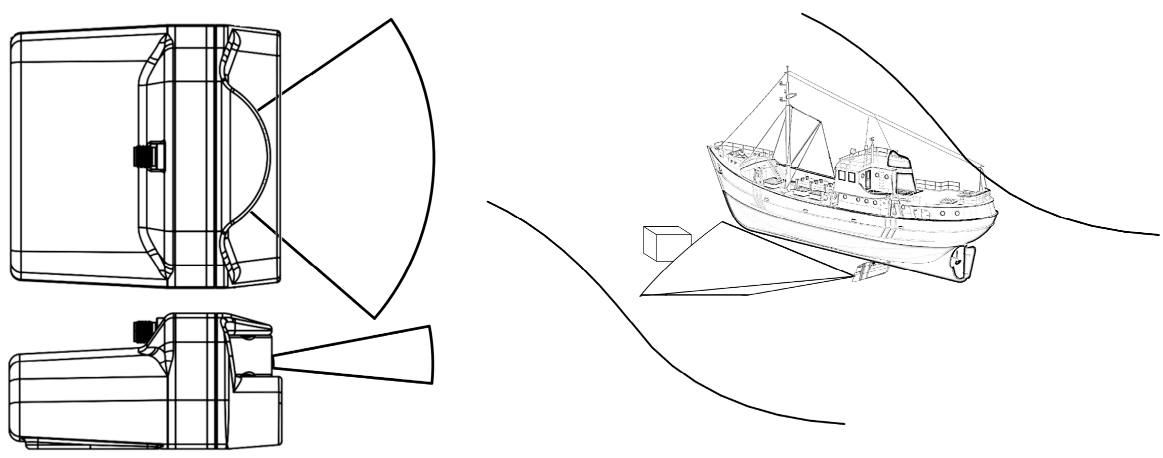
This paper is structured into five sections. First, it introduces the research background and related studies. Subsequently, a method for artificial fish reef posture recognition is proposed, along with a detailed explanation of the dataset construction process. Then, discussions on the experimental improvements based on YOLOv8n-pose are conducted, proposing two enhanced algorithms, namely, YOLOv8n-RFCA-SEGNEXT and YOLOv8 N-MONET-RFCA, based on precision requirements and computational constraints, followed by ablation experiments. Furthermore, a comprehensive discussion of the research findings is presented, concluding with a summary and future research directions, exploring the potential applications of deep learning in artificial fish reef sonar image posture recognition.

2. Materials and Methods

2.1. Data Collection for Artificial Reefs

The study used Oculus sonar to gather data on artificial fish reefs. Oculus is a multibeam sonar system consisting of an array of receivers that collect echoes from a single transmitted pulse and synthesize sonar images through mathematical transformations. The Oculus sonar can produce multiple images per second and allows real-time viewing similar to a camera's output.

The Oculus sonar system is capable of detecting more than one target ahead. As sound pulses reflect off different targets, they return to the receiver at varying time intervals. The receiver records all these echo signals, covering a time span approximately twice the maximum operational range. These echoes are merged into a single received signal, with their positions on the time axis proportional to the distance of the targets and the signal intensity dependent on the reflectivity of the target surfaces. Targets with different densities, such as gas/air or rock/concrete, exhibit distinct reflection characteristics, resulting in echo signals of varying intensities. For instance, materials like concrete produce strong echoes, while materials like mud, silt, sand, and vegetation produce weaker echo signals [21]. Therefore, the sonar system's acoustic "illumination" range for targets made of concrete is limited, only displaying targets within the sonar beam, as shown in Figure 1a. Targets outside the beam will not appear on the sonar display.



(a) (b)

Figure 1. Figure 1 illustrates the data collection setup involving the boat and sonar. (**a**) presents a schematic diagram of the Oculus sonar product, and (**b**) depicts the installation of the sonar on the boat along with its operational concept.

We installed the Oculus sonar device on the bottom of the boat at a tilt angle of 45 degrees, as illustrated in Figure 1b. As the vessel moved forward, the sonar device scanned the seafloor images. The data were collected from a nearshore area with a water depth of approximately 20 m. The Oculus ViewPoint software v1.21.276 provided by Oculus allowed us to view and export the sonar image data as mp4 videos.

2.2. Artificial Reef Attitude Estimation Methods

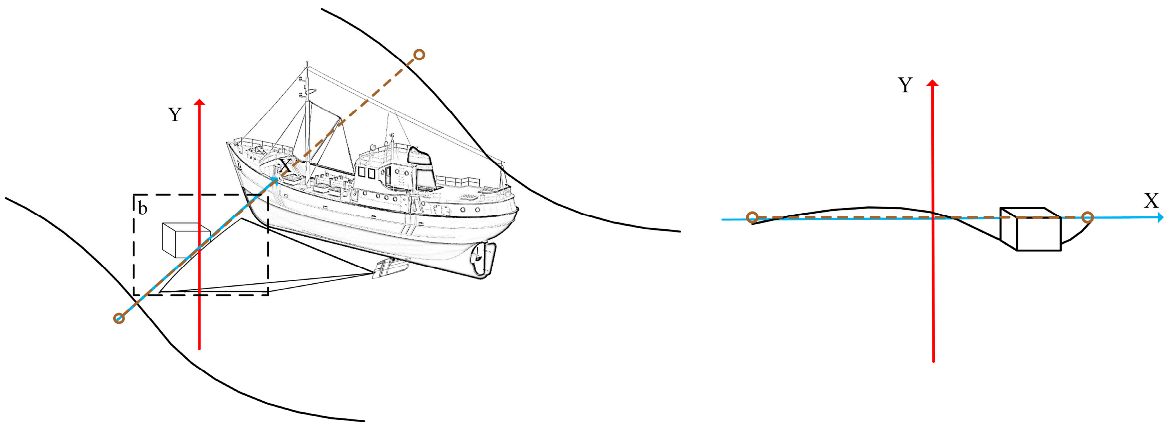
2.2.1. Reef Attitude Estimation Methods

In Table 1, the symbols and definitions related to the mathematical modeling process of attitude estimation for artificial fish reefs are presented.

Table 1. Symbols and definitions.

Symbols	Definitions	Symbols	Definitions
p_1	The first marker of a coral reef.	p_4	The fourth marker of a coral reef.
p_2	The second marker of a coral reef.	$\vec{X}$	The seabed plane.
p_3	The third marker of a coral reef.	$\vec{Y}$	Perpendicular to the seabed plane.

The mathematical modeling of the data collection process involving the horizontal and vertical axes of the ship's direction is presented in this study, as shown in Figure 2a. The perpendicular line on the seabed plane represents the $\vec{Y}$ axis, and the $\vec{X}$ axis is determined by identifying the seabed. The vertical axis, $\vec{Y}$, is constructed in the sonar image perpendicular to the seabed, establishing a two-dimensional Cartesian coordinate system as illustrated in Figure 2b.

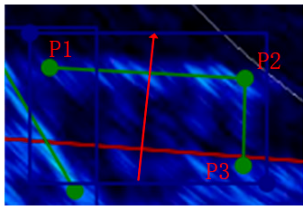
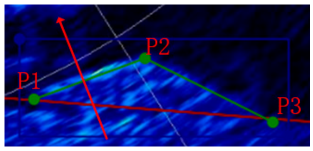
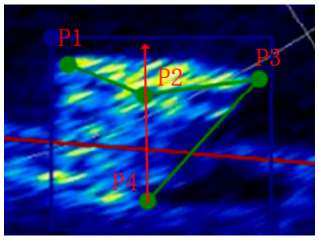


(**a**) (**b**)

Figure 2. The schematic diagram of a three-dimensional coordinate system is shown in Figure 2. (**a**) and (**b**), respectively, illustrate the methods of constructing the coordinate system from different viewing perspectives, with (**b**) being a magnified view of a section of (**a**).

Using this model, we assessed the orientation of artificial reefs. Based on the identification results of different categories, we established the estimation method for the orientation of artificial reefs, as shown in Table 2.

Table 2. Pose estimating method.

Marker	Label	Angle	Attitude Estimation Paradigm
Reef monomer (fuzzy)	Angle	Right angle	
		Acute angle	
Reef monomer (clear)	Reef		

In the field of pose estimation, we have devised distinct recognition methods tailored to different identification scenarios. Specifically, for the identification of ambiguous individual coral reefs, we have categorized the recognition into two scenarios. In the case of relatively unclear coral reef individuals, we first construct the angles identified based on the sequence of points, selecting the perpendicular line of the edge with a larger angle with the $\vec{Y}$ axis as the pose direction. The calculation method is illustrated in Formula (1), where $\overrightarrow{p2p1}_{\perp}$ denotes the normal vector of the vector $\overrightarrow{p2p1}$.

$$\vec{V}_{angle} = \overrightarrow{p2p1}_{\perp} \tag{1}$$

In cases where individual fish reefs can be clearly identified, we use the vector formed by the identified points P4 and P2 to determine the orientation of the fish reef, calculating the angle, as shown in Equation (2). This approach takes into account variations in identification conditions to enhance the accuracy and stability of posture estimation.

$$\vec{V}_{reef} = \overrightarrow{p2p4} \tag{2}$$

The pseudocode implementation of the pose estimation method is shown in Table 3.

Table 3. Pseudocode of the attitude estimation method.

Pseudocode of Attitude Estimation Method
$\vec{Y} = \vec{X}_{\perp}$
If class_label == "angle":
if $\angle\lvert\overrightarrow{p2p1}, \vec{Y}\rvert > \angle\lvert\overrightarrow{p2p3}, \vec{Y}\rvert$:
$\vec{V} = \overrightarrow{p2p1}_{\perp}$
Else:
$\vec{V} = \overrightarrow{p2p3}_{\perp}$
If class_label == "reef":
$\vec{V} = \overrightarrow{p2p4}$

In this study, $\vec{V}$ represents the predicted orientation of artificial reef fish, while $\vec{X}_{\perp}$ denotes the normal vector of vector $\vec{X}$. The angle between vectors $\vec{X}$ and $\vec{Y}$ is shown as $\angle|\vec{X}, \vec{Y}|$.

2.2.2. A Method for Estimating the Subsidence Area of an Artificial Reef

Measuring the settlement area can provide another perspective on the positional and postural changes of artificial reefs. Artificial reefs may experience sinking due to the impact of tidal flushing, and long-term erosion and sedimentation can lead to settlement of the reef body, resulting in a reduction in volume. To detect the settlement areas, the area of settlement can be calculated by determining the region between key points and the seabed plane. The method of assessing the area, as shown in Figure 3, can involve taking the area between the identified settlement region and the seabed plane as the settlement area.

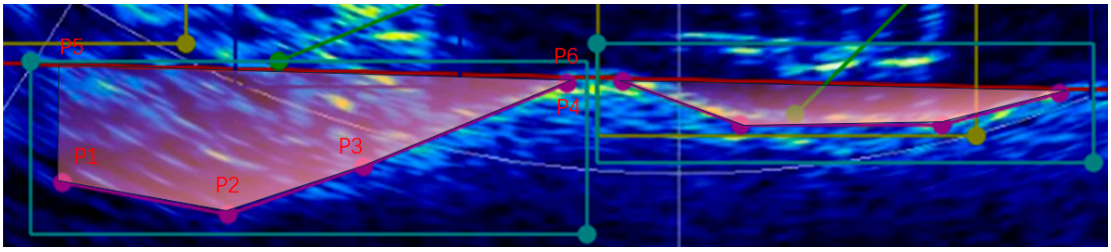


Figure 3. Settling area.

The calculation formula for the settling area in Figure 3 is represented as Formula (3), where p_5 and p_6 refer to the projection points of p_1 and p_4 along the $\vec{Y}$-axis on the $\vec{X}$-axis, respectively.

$$S_{pit} = S_{p_1p_2p_3p_4p_5p_6} \tag{3}$$

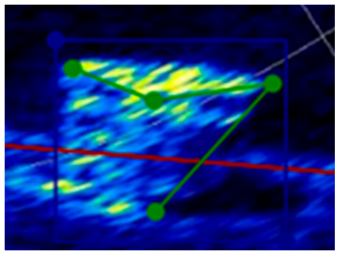
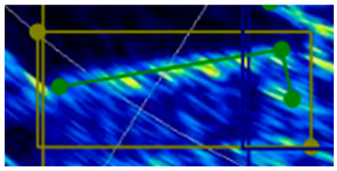
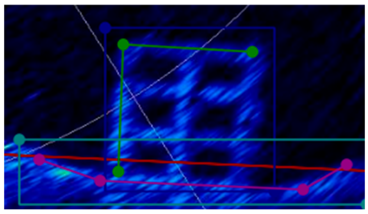
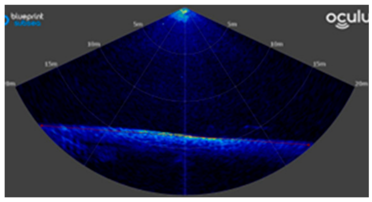
2.3. *Artificial Reef Dataset*

2.3.1. Data Annotation Methods

We have developed an artificial reef dataset based on the pose acquisition method described in Section 2.2. For the annotation task of artificial reefs, we used Labelme 5.3.1 software. The annotation process consists of two steps: object localization and key point annotation. In the object localization step, we employed bounding boxes to mark the position of each reef accurately. The bounding boxes should precisely enclose the entire reef. Subsequently, in the key point annotation step, we annotated the key points following the principles of left to right and top to bottom. Key points were labeled according to the features of the target, leading to four categories of annotation targets.

In Table 4, four categories are annotated with sonar images. The first category consists of clearly visible cube-shaped fish reefs. For these reefs, it is necessary to draw target boxes during the object localization phase and mark key characteristic points during the feature point annotation phase. The second category includes fish reefs, where only one corner of the square is vaguely visible. In these cases, target localization and feature point annotation are also required, with a focus on marking the corner points of the square. The third category comprises concave areas around the fish reefs. These concave areas may be connected to the fish reefs, so it is important to ensure they are included in the target box during the localization phase. The fourth category is the seabed plane. For the seabed plane, a complete image serves as the target identification annotation box to indicate the presence of a seabed plane target in the image. Subsequently, annotation is performed by marking the two ends of the seabed plane. Following these steps, we can accurately annotate different types of targets in artificial fish reef sonar images, as shown in Figure 4.

Table 4. Artificial reef sonar image annotation category.

Label Target	Dimension Label	Explain	Image
Reef monomer (clear)	Reef	You can clearly identify the cube reef	
Reef monomer (fuzzy)	Angle	Only one corner of the reef square can be seen	
Settling area	Pit	Reef sinking due to tidal scouring	
bottom plane	Bottom	underwater plane of a sonar image	

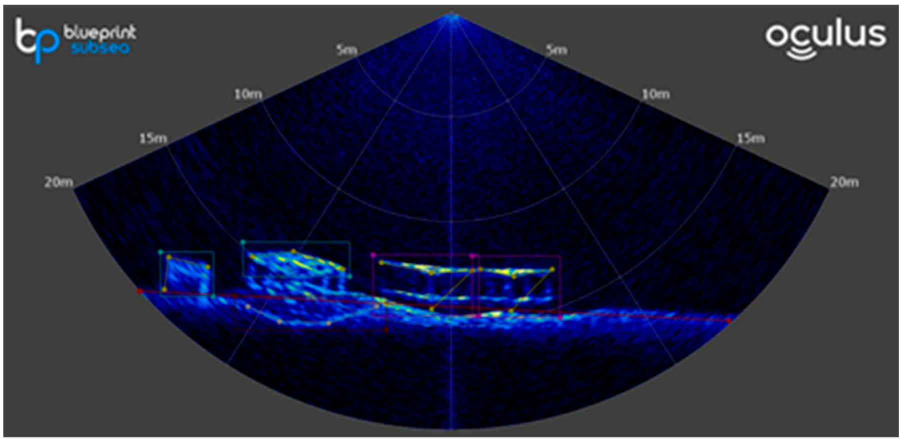


Figure 4. Complete annotation results of artificial reef.

A subset of the collected videos was processed to extract frames, resulting in 2630 images. These images were randomly divided into training, testing, and validation sets at an 8:1:1 ratio. Following the division, the training set comprised 2103 images, the validation set contained 263 images, and the testing set included 264 images. Such data partitioning aids in ensuring that the model accounts for various sample scenarios during training, validation, and testing phases, thereby enhancing the model's generalization ability and accuracy. The target quantities in each dataset after partitioning are presented in Table 5.

Table 5. Number of targets in the dataset.

Dataset	Bottom	Angle	Reef	Pit
Train	2103	3873	1415	483
Val	263	452	190	69
Test	264	542	203	62

2.3.2. Data Enhancement Methods

In the process of model training, we used the data augmentation methods integrated within the YOLO v8 framework to randomly enhance the pre-segmented dataset. These data augmentation techniques include Mosaic, Random Affine, and MixUp, which play a crucial role in enhancing the model's understanding of scenes and its ability to learn details. Particularly, the Mosaic method, derived from the YOLO v4 framework [22], can concatenate multiple images to form a single "Mosaic" image. This type of data augmentation helps enrich the training data and improve the model's generalization capability. In each batch, we concurrently applied various data augmentation techniques such as random cropping, composition, and rotation to further enhance the diversity and complexity of the training images. By transforming original images into suitable training images through data augmentation methods, as shown in Figure 5, the model in the training process will learn more features and patterns under different circumstances, thereby enhancing its performance in various application scenarios.

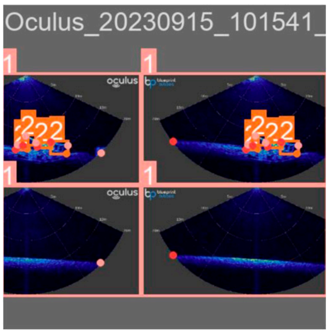


Figure 5. Image after data enhancement.

After data augmentation, the training images provide the model with a greater variety of samples to learn from, enabling it to better tackle various data changes and details. This training approach effectively enhances the model's generalization ability and adaptability, leading to more robust and efficient performance in practical applications.

2.4. Estimation of Artificial Reef Pose

In the process of estimating the posture of artificial reefs, the first step involves processing the sonar images obtained and using the keypoint detection model to identify the seabed plane, individual reef structures, and their corresponding key points. This initial procedure aids in determining the position and shape of the reefs, laying the groundwork for subsequent posture estimations. Subsequently, a two-dimensional coordinate system is established based on the identified seabed plane, enabling a more accurate description of the reefs' positions and orientations. By employing this coordinate system in conjunction with the key point information of the reef structures, the posture of the reefs can be calculated using the estimation methods provided in Table 2. The crux of the entire process lies in transforming sonar images into coordinate information, followed by posture estimation integrating target recognition and keypoint extraction results. Through this process, a more precise understanding of the morphological characteristics of artificial reefs can be achieved, offering vital data support for subsequent research and analysis. The specific implementation process is illustrated in Figure 6.

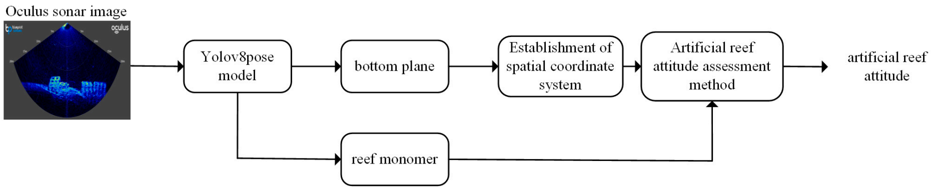


Figure 6. The process diagram for estimating the posture of artificial fish reefs.

2.4.1. Key Point Recognition Method Based on Deep Learning

Recently, there has been rapid development in deep learning technology in the field of image processing, especially in tasks such as image classification, object detection, semantic segmentation, and instance segmentation. Key point detection, as an extension of object detection, involves identifying key points within the detected objects, enabling machines to gain a deeper understanding of the structure and boundary features of the detected objects. Deep learning models for key point detection recognize key points within the detected object area, aiding machines in comprehensively grasping the content of the image. The process of recognition is illustrated in Figure 7.

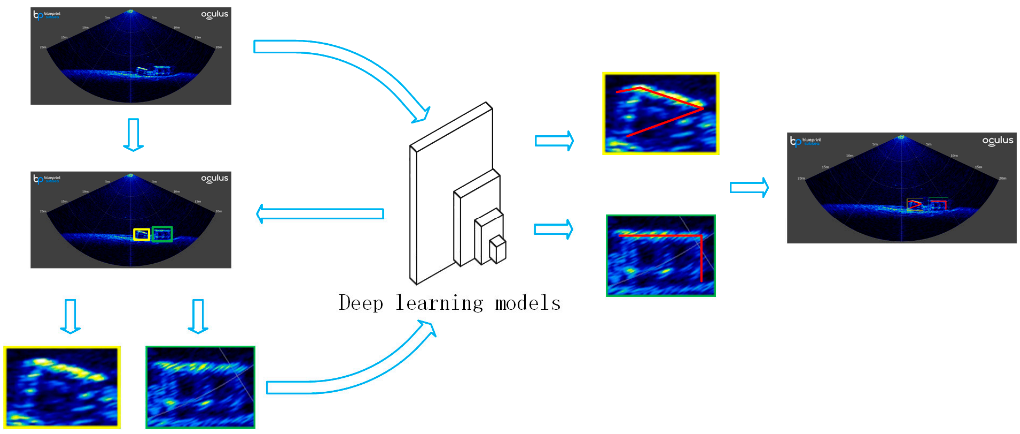


Figure 7. Key point recognition method.

In practical applications, deep learning models are trained by learning from a large volume of image data to precisely detect key points within images. The model's workflow typically encompasses data preprocessing, feature extraction, model training, and inference steps. By continuously optimizing both the model structure and parameters, contemporary deep learning algorithms have shown outstanding performance in image processing tasks, playing a vital role in diverse practical applications.

Based on the artificial reef sonar image dataset established previously, this study employs the YOLOv8-pose model within the YOLOv8 framework to accomplish the task of artificial reef posture recognition. The choice of the YOLOv8n-pose model is based on its robust engineering framework in the field of object recognition and its high efficiency on edge devices. The YOLOv8 series models are recognized for their exceptional performance and straightforward deployment, with YOLOv8n being particularly popular for seamless integration on edge devices. This facilitates the estimation of artificial reef poses during navigation using sonar. Additionally, significant research has been dedicated to optimizing operators and exploring practical deployment scenarios of YOLO series algorithms on edge devices. Therefore, YOLOv8npose has been selected as the model for key point recognition. Although not the absolute optimal choice, it stands out as one of the most practical models for engineering applications and deployment. YOLOv8npose, an extension of the YOLO series models, has been widely applied across various domains and has undergone numerous optimization enhancements by researchers, specifically tailored for key point recognition [23–27].

2.4.2. YOLOv8-Pose Model Structure

The YOLOv8-Pose model is a keypoint detection model based on the YOLO object detection algorithm. This model combines the real-time capabilities and efficiency of the YOLO algorithm, enabling rapid and accurate detection of keypoints in images or videos. It has demonstrated high performance in human pose detection, providing reliable technical support for real-time pose analysis. In this study, the YOLOv8-Pose model is applied to the task of detecting artificial reef sonar images. Figure 8 illustrates the standard structure of the YOLOv8-Pose model.

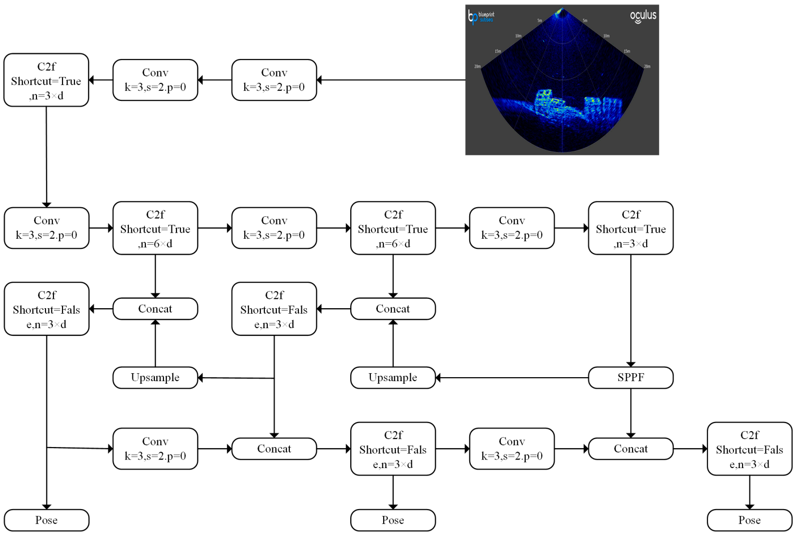


Figure 8. YOLOv8-Pose standard model.

Considering the limited computing capabilities of mobile devices in the marine fishing vessel environment, we opted to experiment and enhance the YOLOv8n-Pose model, which has a small size of only 6.2 M and requires a computational power of just 8.4 GFLOPs. Figure 9 illustrates the results of the artificial reef posture estimation.

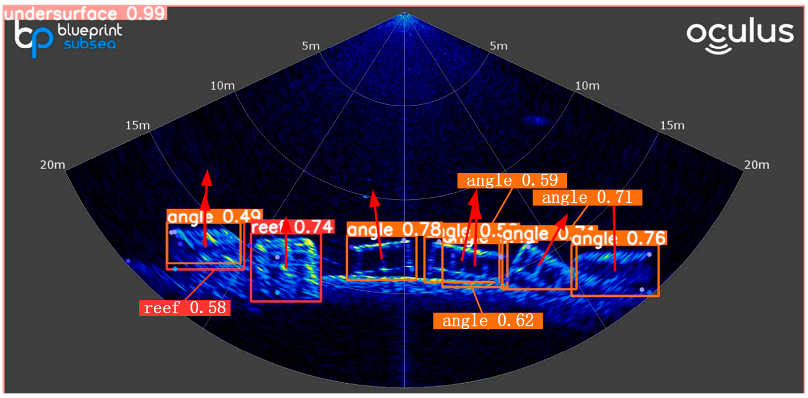


Figure 9. Artificial reef posture estimation results.

2.5. *YOLOv8n-Pose Model Improvement Method*

2.5.1. Accuracy Improvement

In 2022, Guo et al. introduced SegNext, emphasizing the superiority of convolutional attention as a more effective and efficient context encoding mechanism in the field of semantic segmentation compared to the self-attention mechanism in transformers [27]. The structure of SegNext Attention is illustrated in Figure 10. Building on this insight, they developed a novel convolutional attention network architecture called SegNext Attention. As the colors in sonar images are primarily uniform and the features capture the intensity variations of the returned sound waves, highlighting the boundaries between different objects in artificial reef sonar images is crucial. Therefore, the integration of SegNext Attention in the field of semantic segmentation is expected to significantly enhance the model's ability to accurately identify artificial reef targets.

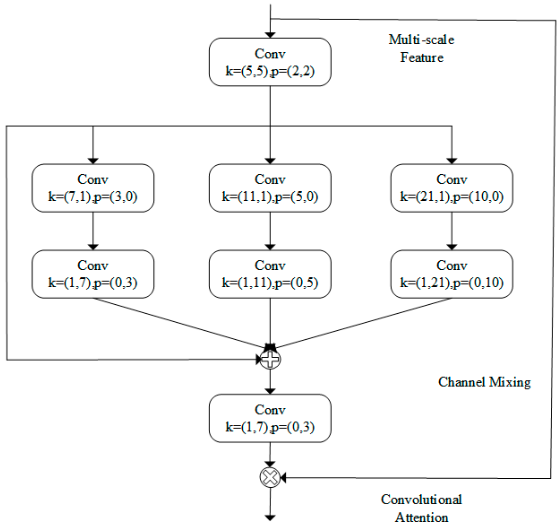


Figure 10. SegNext_Attention structure diagram.

In 2023, Zhang et al. introduced a novel attention mechanism known as Receptive Field Channel Attention (RFCA), which enhances the convolutional structure RFCAConv as illustrated in Figure 11. RFCA focuses not only on spatial features within the receptive field but also provides effective attention weights for large convolutional kernels [28]. YOLOv8n, a prominent model in object detection, follows a one-stage detection approach that combines fast inference speed with high accuracy. By integrating the RFCAConv structure, YOLOv8n can access richer contextual and spatial information, thereby improving the model's perceptual capability and detection accuracy.

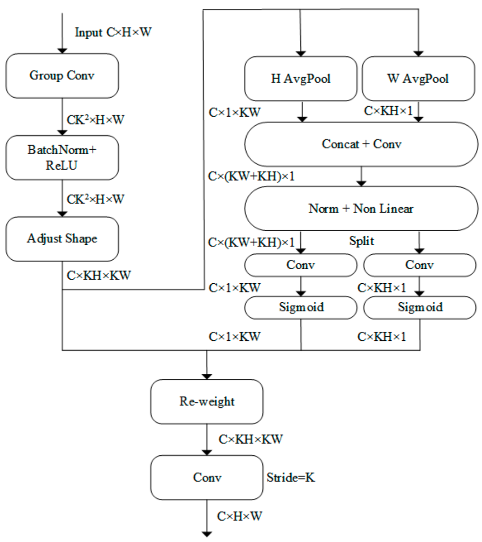


Figure 11. RFCAConv structure diagram.

The improvement of the YOLOv8n-Pose model was achieved by integrating the strengths of Guo et al.'s SegNext Attention and Zhang et al.'s RFCAConv algorithms. By incorporating the convolutional attention mechanism of SegNext Attention, the model can better encode contextual information, enhancing the accuracy of artificial reef recognition.

Additionally, the integration of RFCAConv's receptive field attention mechanism effectively addresses the challenge of shared convolutional kernel parameters, providing more accurate attention weights for large kernel sizes. This combination brings richer spatial and contextual information to the model, further enhancing its perceptual capabilities and detection accuracy in object detection tasks. By combining these advanced structures with the YOLOv8n-Pose model, the enhanced model is expected to achieve significant improvements in performance and accuracy on the artificial reef dataset. We introduce our model as the YOLOv8n-PoseRFSA, where 'RF' stands for RFCAConv structure and 'SA' stands for SegNext Attention structure. The architectural diagram of this model is illustrated in Figure 12.

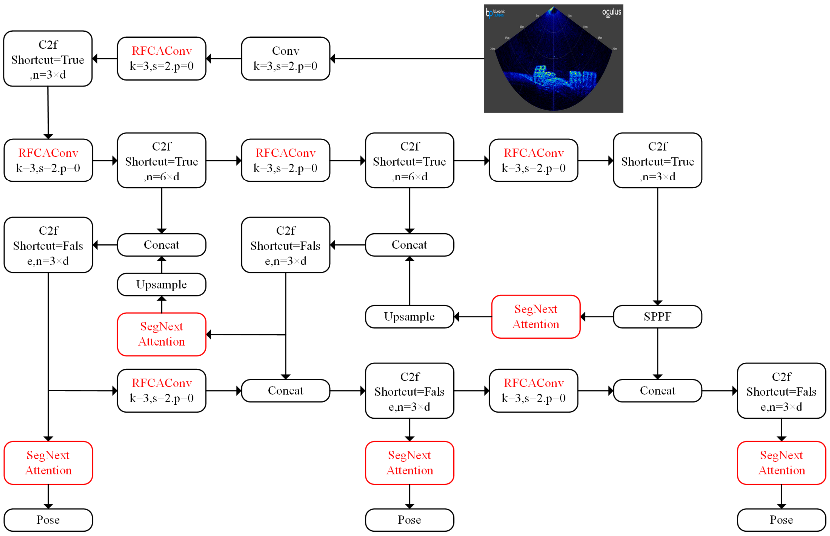


Figure 12. YOLOv8n-PoseRFSA structure diagram.

2.5.2. Miniaturization Improvement

MobileNetV3 [29] is a lightweight neural network architecture particularly well-suited for real-time object detection and recognition tasks on mobile and embedded devices, as illustrated in Figure 13. Using MobileNetV3 as the backbone network for YOLOv8 enables a faster and lighter object detection algorithm, which has been proven in many enhanced models of YOLOv8. MobileNetV3 reduces model parameters and computational complexity while maintaining high detection accuracy, making the enhanced YOLOv8n-PoseRFSA model suitable for deployment on mobile or embedded devices.

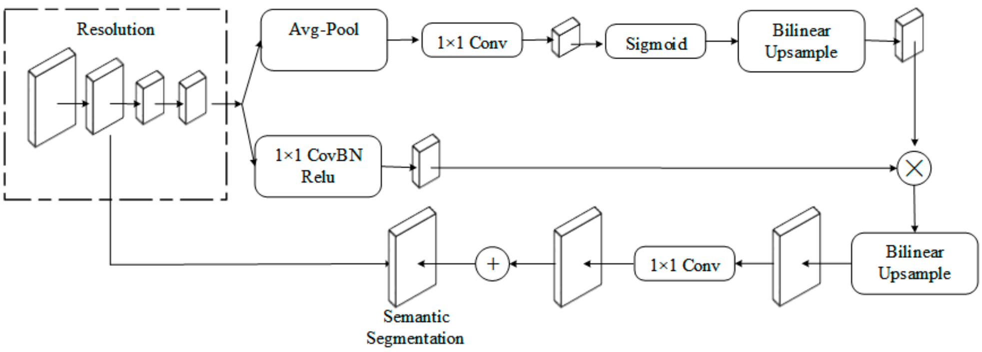


Figure 13. MobileNetV3 structure diagram.

The replacement of the YOLOv8n-PoseRFSA backbone network with the MobileNetV3 model is depicted as the YOLOv8n-PoseMRFSA model in Figure 14, where 'M' stands for MobileNetV3 structure.

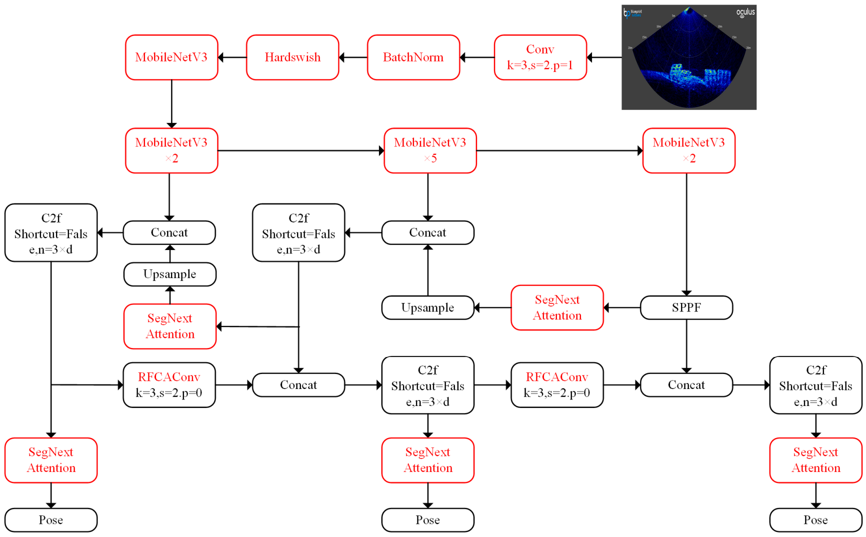


Figure 14. YOLOv8n-PoseMRFSA structure diagram.

In the YOLOv8n-PoseMRFSA network, MobileNetV3 demonstrates excellent efficacy in feature extraction. However, when integrated into the YOLOv8n-PoseRFSA model, specifically designed for artificial reef recognition, it is crucial to evaluate MobileNetV3's influence on the feature extraction process. Therefore, conducting comprehensive ablation experiments is essential to enhance the model and mitigate any potential decrease in the accuracy of the YOLOv8n-PoseRFSA model resulting from potential changes in feature extraction structures.

3. Results

3.1. Ablation Experiments

3.1.1. Evaluation Indicators

In the field of object detection, commonly used evaluation metrics include the following:

1. Precision (P): Precision is defined as the number of true positive predictions made by the model divided by the total number of positive predictions made by the model. It measures the accuracy of the model in predicting positives. Here, TP represents true positives, and FP represents false positives.

$$P = \frac{TP}{TP + FP} \tag{4}$$

2. Recall (R) is defined as the number of true positive instances correctly predicted by the model divided by the total number of actual positive instances in the dataset. It quantifies the model's ability to correctly identify all relevant instances. Here, FN represents false negatives.

$$R = \frac{TP}{TP + FN} \tag{5}$$

3. Metric/mAP50, short for mean average precision at 50, is a commonly used evaluation metric in object detection tasks such as YOLOv8. It calculates the average precision of a model at various confidence levels, considering only the top 50% of detection results with the highest confidence scores. Here, N represents the number of classes, and AP_i^{50} denotes the average precision of the i-th class on the top 50 detection results.

$$\text{metrics/mAP50} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i^{50} \tag{6}$$

4. The metrics/mAP50-95, which stands for the mean average precision at different IoU thresholds ranging from 0.5 to 0.95, are another evaluation measure for object detection models. It computes the average precision of the model across a range of IoU thresholds, providing a comprehensive assessment of the model's performance at various levels of intersection between predicted and ground truth bounding boxes. Here, AP_i^{50-95} represents the average precision for the i-th class within the threshold range of 50 to 95.

$$\text{metrics/mAP50-95} = \frac{1}{N}\sum_{i=1}^{N} AP_i^{50-95} \tag{7}$$

5. GFLOPs (giga floating point operations per second) is a commonly used metric for assessing the computational load of a model. It indicates the number of billion floating-point operations performed per second and is used to evaluate the computational complexity and efficiency of a model. When evaluating model performance, GFLOPs are frequently employed to measure the computational resources required during the training or inference stage, aiding in the selection of appropriate hardware devices or optimizing model structures.

3.1.2. Ablation Experimental Results

Experimentation through ablation is a commonly used approach in research aimed at elucidating the significance and functionalities of individual components or features within deep neural network models, with the ultimate goal of enhancing model performance or optimizing model architecture. In this study, three models were compared: YOLOv8n-Pose, YOLOv8n-PoseRFSA, and YOLOv8n-PoseMRFSA, and ablation experiments were conducted on each enhanced model. The software environment comprised the Python 3.9 programming language, the PyTorch 2.0.0 deep learning framework, the CUDA 11.4 operating platform, and Anaconda 23.1.0 environment management software, all operating on the Ubuntu 18.04.5 LTS system with a Tesla V100S 32G graphics card. During model training, all images were resized to 640 × 640 dimensions using the AdamW optimizer with lr = 0.00125 and momentum = 0.9, and the maximum number of epochs was set to 400. Results from the ablation experiments can be found in Table 6.

Table 6. Ablation experiment results.

Model	Box				Pose				GFLOPs
	P	R	mAP50	mAP50-95	P	R	mAP50	mAP50-95	
YOLOv8n-Pose	0.886	0.868	0.926	0.712	0.871	0.856	0.907	0.839	8.4
YOLOv8n-PoseSA	0.906	0.89	**0.949**	0.715	0.891	0.876	**0.926**	0.86	9.1
YOLOv8n-PoseRF	0.903	0.898	0.942	0.715	**0.895**	**0.889**	0.924	**0.865**	8.7
YOLOv8n-PoseRFSA	**0.909**	**0.904**	0.944	**0.72**	0.892	0.888	0.92	0.857	9.2
YOLOv8n-PoseM	0.868	0.855	0.918	0.675	0.862	0.847	0.895	0.792	2.8
YOLOv8n-PoseMRF	0.867	0.884	0.919	0.676	0.842	0.881	0.901	0.786	2.8
YOLOv8n-PoseMSA	0.918	0.818	0.924	0.689	0.908	0.809	0.901	0.793	3
YOLOv8n-PoseMRFSA	0.854	0.853	0.899	0.653	0.834	0.834	0.882	0.751	3.1

The bolded content represents the best result for this comparison.

Through ablative experiments, we systematically removed specific components from the model to investigate their impact on overall performance, thereby gaining a deeper understanding of the model structure and the contributions of individual components. According to the experimental results in Table 5, the YOLOv8n-PoseRFSA model showed improvements in various evaluation metrics compared to the YOLOv8n-Pose model, with an approximately 10% increase in computational load. In this model, the precision of the box increased by 0.023, the recall by 0.036, and both Map50 and Map75 improved by 0.018 and 0.008, respectively. In terms of pose detection, precision increased by 0.021, recall by 0.032, and Map50 and Map75 by 0.013 and 0.018, respectively. Moreover, in the improved YOLOv8n-PoseMSA model, the performance of the YOLOv8n-PoseMSA model stood out. Although there was a slight decrease in the Map metric for this model, both precision and recall showed improvements. Most importantly, the computational load of the YOLOv8n-PoseMSA model was only 33% of the YOLOv8n-Pose model, aligning with the optimization goal of significantly reducing model computation while accepting a slight decrease in accuracy.

3.2. Model Validation

To evaluate the practical recognition performance of the model, this study conducted a comparative verification experiment between the model's recognition results and manual counting. The data for the verification experiment were obtained from sonar videos of artificial reefs used to construct the dataset. Twenty images with varying numbers of reef structures were randomly selected from the images containing reefs. In the comparative verification experiment, manual counting, the YOLOv8n-Pose model, the YOLOv8n-PoseRFSA model, and the YOLOv8n-PoseMRFSA model were used to count the reef structures in these 20 sonar images of artificial reefs. The recognition parameters were set to conf = 0.25 and iou = 0.7, following the default parameters of the YOLOv8n-Pose model. The target reef structures were identified as the sum of the detected reference and proxy markers, with the results of the verification experiment presented in Table 7.

Table 7. Validation of experimental results.

Number\Reef Target	Manual Counting	YOLOv8n-PoseRFSA	YOLOv8n-PoseMSA	YOLOv8n	Number\Reef Target	Manual Counting	YOLOv8n-PoseRFSA	YOLOv8n-PoseMSA	YOLOv8n
1	1	1	1	1	11	3	4	4	3
2	5	5	5	7	12	1	1	1	1
3	4	5	4	4	13	3	3	3	2
4	3	3	3	3	14	6	5	6	7
5	5	5	9	7	15	6	5	6	5
6	4	3	6	2	16	4	6	5	5
7	8	7	8	6	17	2	3	2	3
8	6	6	7	6	18	3	5	4	4
9	8	9	10	6	19	4	5	6	3
10	6	5	5	5	20	1	1	1	1

In order to more specifically measure the differences between different models, variance was employed to compare YOLOv8n-pose with the improved model against manually recognized results. The statistical findings are presented in Table 8.

Table 8. Statistical results.

Model	Variance
YOLOv8n-Pose	0.81
YOLOv8n-PoseMSA	0.5625
YOLOv8n-PoseRFSA	0.49

The variance between the YOLOv8n-PoseRFSA model and the manual counting results was minimized at the highest level by YOLOv8n-Pose, a finding that aligns with the outcomes of the ablation experiments.

4. Discussion

4.1. Dataset Acquisition

The dataset of artificial reefs collected in this study was obtained through scanning and image capture of the seabed with the Oculus sonar system installed on the bottom of the vessel. The sonar system is capable of receiving echo signals from targets, thereby displaying the distances and features of different targets. In contrast to the side-scan sonar dataset used by Xiong et al. in 2014, the data used in this study were acquired from an imaging sonar, allowing for a more accurate determination of the orientation of artificial reefs. Within the detection range of sonar, it is possible to distinguish artificial reefs from other seabed environments relatively accurately. Compared to using cameras to capture underwater environments, sonar images have the advantage of distinguishing different materials, better showcasing the material characteristics of underwater objects, and not being restricted by the light sources used in camera shots.

During the data annotation process, challenges were faced when annotating artificial reefs with significant overlap, where only a fragment of the square shape of the reef is visible due to the overlapping. This poses challenges as the human eye can only identify the top layer, making it relatively easy to annotate reefs with distinct square features. In cases where there is significant overlap of complete reef structures, sonar images may struggle to clearly differentiate them.

4.2. Attitude Estimation Method for Artificial Reefs

The method for estimating the posture of artificial reefs based on sonar images is constructed using YOLOv8n-Pose recognition results. While pose recognition introduced by YOLOv8 is primarily used for identifying human key points, in this study, it has been successfully applied to identifying key points of artificial reefs. By successfully identifying the key points of artificial reefs, we have established a method for estimating the posture of artificial reefs. It should be noted, however, that this method is based on 2D images generated by sonar imaging. As a result, it can only, to some extent, accomplish the task of estimating the posture of artificial reefs. In comparison to previous assessments of artificial reefs through side-scan sonar, which only identified the quantity of reefs and the approximate height of reef clusters, the posture estimation method developed in this paper can better track the posture, laying a foundation for more detailed research on artificial reefs.

However, compared to completing a full 3D modeling of artificial reefs for posture estimation, this method has certain limitations in terms of the posture data it provides, as it can only estimate the angles of artificial reefs. Although it is comparatively challenging to conduct comprehensive 3D modeling of reefs underwater, given that artificial reefs typically have regular shapes, our method can accomplish the task of estimating the posture of artificial reefs under relatively simple conditions. Further research and discussion are needed on how to combine more advanced image processing technologies and algorithms to enhance the accuracy and comprehensiveness of posture estimation on artificial reefs. In recent years, numerous scholars have emphasized the importance of artificial coral reefs in ocean ecosystem restoration and biodiversity enhancement. Therefore, exploring how to integrate sonar images and other sensor data to achieve a more comprehensive and accurate estimation of artificial reef postures will be one of the directions for future research. Through continuous optimization and improvement of methods, we can better understand and monitor artificial reefs, providing more support and data for the conservation and management of marine ecosystems.

4.3. Experimental Results

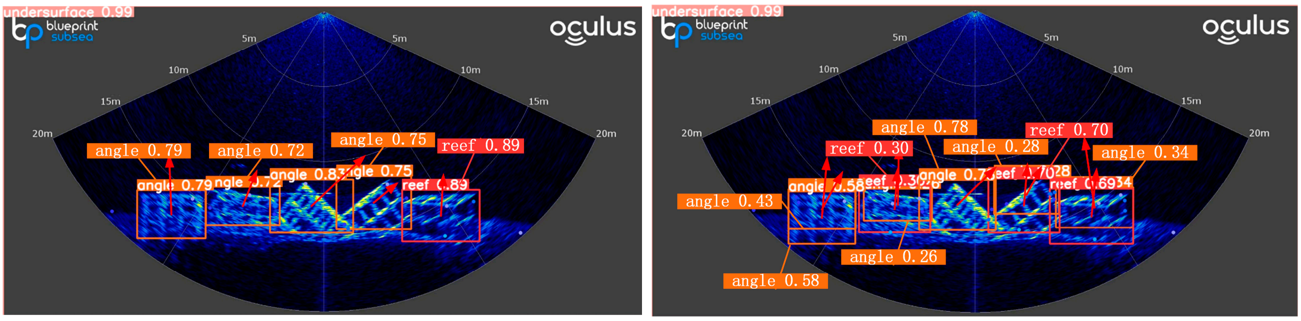
4.3.1. Ablation Experiments

In this ablative experiment, we conducted an in-depth study and evaluation of the refined model by systematically removing specific components from the model. Among them, the YOLOv8n-PoseRFSA model and the YOLOv8n-PoseMRFSA model performed exceptionally well. The YOLOv8n-PoseRFSA model combines the RFCAConv and SegNext_Attention modules, resulting in improvements in various evaluation metrics compared to the base YOLOv8n model, despite an increase of approximately 10% in computational complexity. This model achieved significant performance gains in both object detection and pose estimation tasks, manifesting improvements in precision, recall, and various Map metrics. The synergistic integration of the RFCAConv and SegNext_Attention modules in this composite model, as opposed to the introduction of individual modules, had a more pronounced effect, playing a pivotal role in enhancing the overall model performance.

On the other hand, the YOLOv8n-PoseMSA model incorporates the RFCAConv module into the MobileNetV3 backbone network. Although this model exhibited a slight decrease in Map metrics, it demonstrated notable improvements in precision and recall. Notably, the computational complexity of this model is only one-third of the base YOLOv8n model, significantly reducing computational costs while maintaining a certain level of accuracy, aligning with the optimization goal of balancing accuracy and efficiency. Overall, the YOLOv8n-PoseRFSA and YOLOv8n-PoseMSA models showcased favorable performance in this experiment. The successful validation of these refined models emphasizes the significant impact of incorporating new perceptual modules, attention mechanisms, and integrating lightweight network backbones on model performance.

4.3.2. Validation Experiments

In most cases, the difference in recognition performance between the two models in Table 7 is minimal, with an error margin of only one to two coral colonies compared to manual identification. Therefore, these models can effectively substitute human effort for accurately counting artificial coral reefs. However, there are certain differences in the recognition process between the YOLOv8n-PoseRFSA model and the YOLOv8n-PoseMSA model. One of the most noticeable examples is image No. 5 in Table 7, where the recognition results using different models for image No. 5 are shown in Figure 15.



(**a**) (**b**)

Figure 15. Recognition results of image No. 5 on different models. (**a**) displays the pose estimation results for image No. 5 using the YOLOv8n-PoseRFSA model, and (**b**) shows the pose estimation results for image No. 5 using the YOLOv8n-PoseMSA model. The arrow indicates the direction of the artificial reef posture identified by the method.

In the recognition of the objects in image No. 5, the YOLOv8n-PoseRFSA model produced results consistent with manual counting, while the YOLOv8n-PoseMSA model showed a higher frequency of duplicate recognitions, indicating its lower robustness. In terms of pose estimation methods, it can be observed from image No. 5 that, when correctly identifying the coral reefs, the estimated poses by the method closely align with human judgment. However, for the coral reefs erroneously recognized by the YOLOv8n-PoseMSA model, significant orientation discrepancies are evident. This may arise from the standard cuboid composition of coral reefs, where the model might mistake the internal framework of artificial reefs for the reef outline.

5. Conclusions

This paper focuses on artificial reef posture recognition, introducing a novel method based on sonar imaging to determine the posture of artificial reefs. Initially, an artificial reef dataset was constructed using Oculus sonar, and the YOLOv8n-Pose framework was employed in model design. Various model enhancement techniques were implemented to increase accuracy and efficiency. The method for artificial reef posture recognition was proposed, detailing the dataset construction process, discussing experiments on YOLOv8n-Pose enhancements, and presenting two improvement algorithms. By exploring a novel approach to artificial reef posture recognition, the study aims to advance technology and address pertinent issues. The proposed method for automated detection of artificial reefs in this study enables real-time monitoring of deployment status and quantification of their posture information, significantly enhancing monitoring efficiency and convenience. Notably, this method allows for a more precise assessment of the layout of artificial reefs and their impact on marine ecosystems, providing additional data support for future reef design and deployment.

6. Outlook

The proposed method for artificial reef attitude estimation in this paper uses sonar images to estimate the attitude of artificial reefs on the seabed. However, improvements are needed in various aspects, including data collection, estimation methods, data processing, and visualization.

1. Addressing the challenge of overlapping angles of artificial reefs on the seabed that may go undetected by a single sonar system, future research could focus on exploring the principles of 3D modeling and reconstruction by integrating data from multiple sonar perspectives, similar to the methodologies employed in radar for 3D modeling.
2. The latest YOLOv9 model on the COCO dataset for object detection demands 26.4 GFLOPs, significantly higher than that of YOLOv8n (8.7 GFLOPs). Despite the absence of a YOLOv9 model with computational requirements akin to YOLOv8n or the release of YOLOv9s in this paper and source code, this study abstained from experimenting with YOLOv9. Nonetheless, given the comparable performance of YOLOv9 at a similar scale, it showcases extensive potential applications.
3. When using sonar for artificial reef detection, it is advantageous to integrate the examination of fish and other organisms' distribution within the artificial reef deployment zones. This necessitates simultaneous data collection and analysis in conjunction with optical devices to effectively enhance the conservation and administration of marine ecosystems.

Author Contributions: Y.S. (Yifan Song): conceptualization, methodology, software, investigation, writing—original draft, writing—review and editing. Z.W.: resources, data curation, writing—review and editing. S.Z.: conceptualization, methodology, writing—review and editing. W.Q.: investigation, resources, supervision. Y.S. (Yongchuang Shi): formal analysis, investigation. X.X.: visualization, formal analysis. P.L.: visualization, formal analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (32201298), the Central Public-Interest Scientific Institution Basal Research Fund (2021T01), and the Technology Development Projects related to the safety of cold sources in nuclear power plants (21FW018).

Data Availability Statement: We recognize the importance of data and understand that other researchers may be interested in our study and would be happy to provide further support and assistance if they require data support or would like to communicate with us. We can be contacted by email or other appropriate means and will be happy to support you with data or related research.

Acknowledgments: The study was conducted in the waters of Linhai City, Zhejiang Province. This research recognizes the generous support from the National Natural Science Foundation of China and the Central Public-Interest Scientific Institution Basal Research Fund.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Higgins, E.; Metaxas, A.; Scheibling, R.E. A systematic review of artificial reefs as platforms for coral reef research and conservation. *PLoS ONE* **2022**, *17*, e0261964. [CrossRef] [PubMed]
2. da Silva, G.V.; Hamilton, D.; Murray, T.; Strauss, D.; Shaeri, S.; Faivre, G.; da Silva, A.P.; Tomlinson, R. Impacts of a multi-purpose artificial reef on hydrodynamics, waves and long-term beach morphology. *J. Coast. Res.* **2020**, *95*, 706–710. [CrossRef]

3. Glarou, M.; Zrust, M.; Svendsen, J.C. Using artificial-reef knowledge to enhance the ecological function of offshore wind turbine foundations: Implications for fish abundance and diversity. *J. Mar. Sci. Eng.* **2020**, *8*, 332. [CrossRef]
4. Firth, L.B.; Farnworth, M.; Fraser, K.P.P.; McQuatters-Gollop, A. Make a difference: Choose artificial reefs over natural reefs to compensate for the environmental impacts of dive tourism. *Sci. Total Environ.* **2023**, *901*, 165488. [CrossRef] [PubMed]
5. Acarli, D.; Kale, S. Species composition of artificial reef models specifically designed for Homarus gammarus (Crustacea: Decapoda: Nephropidae) in the Sea of Marmara. *Res. Mar. Sci.* **2020**, *5*, 625–635.
6. Zhang, R.; Zhang, H.; Liu, H.; Zhao, J. Differences in trophic structure and trophic pathways between artificial reef and natural reef ecosystems along the coast of the North yellow Sea, China, based on stable isotope analyses. *Ecol. Indic.* **2021**, *125*, 107476. [CrossRef]
7. Guo, Z.; Wang, L.; Cong, W.; Jiang, Z.; Liang, Z. Comparative analysis of the ecological succession of microbial communities on two artificial reef materials. *Microorganisms* **2021**, *9*, 120. [CrossRef] [PubMed]
8. Reis, B.; van der Linden, P.; Pinto, I.S.; Almada, E.; Borges, M.T.; Hall, A.E.; Stafford, R.; Herbert, R.J.; Lobo-Arteaga, J.; Gaudêncio, M.J.; et al. Artificial reefs in the North–East Atlantic area: Present situation, knowledge gaps and future perspectives. *Ocean Coast. Manag.* **2021**, *213*, 105854. [CrossRef]
9. Vivier, B.; Dauvin, J.C.; Navon, M.; Rusig, A.-M.; Mussio, I.; Orvain, F.; Boutouil, M.; Claquin, P. Marine artificial reefs, a meta-analysis of their design, objectives and effectiveness. *Glob. Ecol. Conserv.* **2021**, *27*, e01538. [CrossRef]
10. Lymperaki, M.M.; Hill CE, L.; Hoeksema, B.W. The effects of wave exposure and host cover on coral-associated fauna of a centuries-old artificial reef in the Caribbean. *Ecol. Eng.* **2022**, *176*, 106536. [CrossRef]
11. Brochier, T.; Brehmer, P.; Mbaye, A.; Diop, M.; Watanuki, N.; Terashima, H.; Kaplan, D.; Auger, P. Successful artificial reefs depend on getting the context right due to complex socio-bio-economic interactions. *Sci. Rep.* **2021**, *11*, 16698. [CrossRef] [PubMed]
12. Paxton, A.B.; Steward, D.N.; Harrison, Z.H.; Taylor, J.C. Fitting ecological principles of artificial reefs into the ocean planning puzzle. *Ecosphere* **2022**, *13*, e3924. [CrossRef]
13. Hemery, L.G.; Mackereth, K.F.; Gunn, C.M.; Pablo, E.B. Use of a 360-degree underwater camera to characterize artificial reef and fish aggregating effects around marine energy devices. *J. Mar. Sci. Eng.* **2022**, *10*, 555. [CrossRef]
14. Mahmood, A.; Bennamoun, M.; An, S.; Sohel, F.; Boussaid, F.; Hovey, R.; Kendrick, G.; Fisher, R.B. Automatic annotation of coral reefs using deep learning. In Proceedings of the Oceans 2016 mts/IEEE Monterey, Monterey, CA, USA, 19–23 September 2016; pp. 1–5.
15. Paxton, A.B.; Newton, E.A.; Adler, A.M.; Van Hoeck, R.V.; Iversen, E.S.; Taylor, J.C.; Peterson, C.H.; Silliman, B.R. Artificial habitats host elevated densities of large reef-associated predators. *PLoS ONE* **2020**, *15*, e0237374. [CrossRef]
16. Gonzalez-Rivero, M.; Beijbom, O.; Rodriguez-Ramirez, A.; Bryant, D.E.; Ganase, A.; Gonzalez-Marrero, Y.; Herrera-Reveles, A.; Kennedy, E.V.; Kim, C.J.; Lopez-Marcano, S.; et al. Monitoring of coral reefs using artificial intelligence: A feasible and cost-effective approach. *Remote Sens.* **2020**, *12*, 489. [CrossRef]
17. Xiong, H.; Liu, L.; Lu, Y. Artificial reef detection and recognition based on Faster-RCNN. In Proceedings of the 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 17–19 December 2021; Volume 2, pp. 1181–1184.
18. Sharan, S.; Kininmonth, S.; Mehta, U.V. Automated cnn based coral reef classification using image augmentation and deep learning. *Int. J. Eng. Intell. Syst.* **2021**, *29*, 253–261.
19. King, A.; Bhandarkar, S.M.; Hopkinson, B.M. Deep learning for semantic segmentation of coral reef images using multi-view information. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019; pp. 1–10.
20. Jocher, G.; Chaurasia, A.; Qiu, J. *Ultralytics YOLO*, version 8.0.0; Computer Software; GitHub: San Francisco, CA, USA, 2023. Available online: https://github.com/ultralytics/ultralytics (accessed on 10 January 2024).
21. Xiong, X.; He, R.; Fan, W.; Wu, Z.; Yu, S.; Wang, Z.; Wang, Y.; Dai, Y. Cause Analysis and Preventive Measures against False Bottoms in Echosounder Data. *Appl. Sci.* **2024**, *14*, 2444. [CrossRef]
22. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
23. Shi, J.; Bai, Y.; Zhou, J.; Zhang, B. Multi-Crop Navigation Line Extraction Based on Improved YOLO-v8 and Threshold-DBSCAN under Complex Agricultural Environments. *Agriculture* **2023**, *14*, 45. [CrossRef]
24. Li, Q.; Nie, S. Detection of ear root feature region in heterogeneous images based on pig head posture. *Acad. J. Comput. Inf. Sci.* **2023**, *6*, 20–32.
25. Wu, Q.; Zhang, L. A Real-Time Multi-Task Learning System for Joint Detection of Face, Facial Landmark and Head Pose. *arXiv* **2023**, arXiv:2309.11773.
26. Wang, S.; Zhang, X.; Ma, F.; Li, J.; Huang, Y. Single-Stage Pose Estimation and Joint Angle Extraction Method for Moving Human Body. *Electronics* **2023**, *12*, 4644. [CrossRef]
27. Guo, M.H.; Lu, C.Z.; Hou, Q.; Liu, Z.; Cheng, M.M.; Hu, S.M. Segnext: Rethinking convolutional attention design for semantic segmentation. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 1140–1156.

28. Zhang, X.; Liu, C.; Yang, D.; Song, T.; Ye, Y.; Li, K.; Song, Y. RFAConv: Innovating Spatital Attention and Standard Convolutional Operation. *arXiv* **2023**, arXiv:2304.03198.
29. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.